Combining self-report dietary assessment instruments to reduce the effects of measurement error

Douglas Midthune
National Cancer Institute

In recognition of his internationally renowned contributions to the field of nutrition epidemiology and his commitment to understanding measurement error associated with dietary assessment.

Objectives

Learning objectives

- Understanding how measurement error leads to loss of precision in estimating diet-health associations
- Learning how to combine self-report dietary instruments to regain precision and improve power to detect associations
- Understanding the limitations of such an approach

Objectives

Why combine self-report instruments?

Impact of measurement error

- Measurement error (ME) in self-report dietary assessment instruments leads to:
  - Bias (attenuation) in estimated diet-health associations
  - Loss of precision in estimated associations
    \[ \frac{\text{E}(\alpha_0)}{\text{s.e.}(\alpha_0)} \lt \frac{\text{E}(\alpha_T)}{\text{s.e.}(\alpha_T)} \]
  - Loss of power to detect associations

Presenters and Collaborators

Sharon Kirkpatrick
Series Organizer

Regan Bailey
Laurence Freedman
Douglas Midthune

Dennis Buckman
Patricia Guenther
Amy Subar

Raymond Carroll
Victor Kipnis
Fran Thompson

Kevin Dodd
Susan Krebs-Smith
Janet Tooze

Why combine self-report instruments?

Regression calibration
Comparing study designs
Limitations and other considerations
Summary
Combining self-report dietary assessment instruments to reduce the effects of measurement error

**Impact of measurement error**

- Statistical methods such as regression calibration can correct for bias due to measurement error
- These methods do not typically recover lost precision or power

**Ways to improve precision and power**

- Increase the sample size
- Decrease the measurement error
  - Improve existing dietary instruments
  - Develop new instruments
  - Combine different self-report dietary instruments

**Combining instruments**

- Examples of self-report instruments that could be combined:
  - Food frequency questionnaire (FFQ)
  - 24-hour dietary recall (24HR)
  - Multiple-day food record (FR)
- Each instrument has its own strengths and weakness

**Self-report dietary instruments**

- **FFQ**
  - Cognitive tasks required are more difficult
  - Estimates usual intake
  - Small within-person variation
- **24HR / FR**
  - Less biased for estimating intake
  - Estimate short-term intake
  - Large within-person variation

**Problem:**

- Traditional 24HR and FR are expensive to administer and/or process
- Not practical for use in large cohort studies

**Examples of self-report instruments that could be combined:**

- Food frequency questionnaire (FFQ)
- 24-hour dietary recall (24HR)
- Multiple-day food record (FR)
Recent developments in dietary assessment
- Self-administered automated 24HR, such as the ASA24 (NCI)
- Automated FR, some using mobile phone technology
- Much less expensive than traditional 24HR/FR
- Practical for use in large cohort studies

Regression calibration
- **Regression calibration** corrects estimated diet-health associations for bias due to ME in reported intake
  - (Relatively) simple and intuitive
  - Applicable in many situations (e.g., linear and logistic regression, survival analysis)
  - Often nearly as efficient as maximum likelihood estimation
  - Extends naturally to combine multiple instruments

**Diet-health model:**
\[
\log \left( \frac{\text{Odds}(Y=1)}{\text{Odds}(Y=0)} \right) = \alpha_0 + \beta_0 T + \beta_1 R
\]
Prediction equation: \(E(T | R) = \lambda_0 + \lambda_1 R\)

Regression calibration: replace \(T\) with its **predicted value** \(E(T | R)\) in diet-health model and perform standard analysis

\(E(T | R)\) the **conditional expectation** (mean) of true intake \(T\) given reported intake \(R\)

\(E(T | R)\) the **best** predictor of \(T\) given \(R\)

Assumption:
- \(R\) has "nondifferential error" with respect to disease \(Y\)
- \(R\) provides no information about disease \(Y\) beyond that provided by \(T\)

Under this assumption, regression calibration estimates are (approximately) **unbiased**
**Review of regression calibration**

- Diet-health model:
  \[ \text{Log(Odds}(Y=1)) = \alpha_0 + \alpha_1 R \]
  \[ \text{Predicted value: } E(T | R) = \lambda_0 + \lambda_1 R \]
- Predicted value \( E(T | R) \) provides no more information about true intake than \( R \)
- As a result, regression calibration does **not** recover power lost due to measurement error

**Can RC be made more powerful?**

- If we can improve prediction of true intake, we can increase precision and power

**Enhanced regression calibration**

- Enhanced RC: predict \( T \) using \( E(T | R, C) \), where \( C \) is an additional variable that:
  1. Helps to **predict** true intake, but
  2. **Not** related to health outcome given true intake
     - **Not** a confounder
     - Has nondifferential error
- Requirement 2) crucial: if \( C \) is related to intake, estimated diet-health association will be **biased**
- Additional **self-report** instruments seem to be perfect candidates for enhanced regression calibration

**Example: Enhanced RC**

- Enhanced RC with two 24HR (\( R \)) and FFQ (\( Q \))
  - Assumption: 24HR **unbiased** for true intake
  - Prediction equation:
    \[ E(T | R_1, R_2, Q) = w \times R + (1-w) \times E(T | Q) \]
    - \( R = \text{mean of two 24HR} \)
    - \( w = \text{var}(u) / (\text{var}(u) + \text{var}(e) / 2) \)
    - \( \text{var}(u) = \text{between-person variance in 24HR} \)
    - \( \text{var}(e) = \text{within-person variance in 24HR} \)
  - Parameters estimated in **linear mixed effects model**

**COMPARING DIFFERENT COMBINATIONS OF SELF-REPORT INSTRUMENTS**
In remainder of talk, we compare 3 possible study designs (or dietary assessment strategies) for estimating dietary intake:

- FFQ alone: one FFQ per subject
- 24HR alone: one or more 24HR per subject
- 24HR and FFQ: one FFQ and one or more 24HR per subject


Study designs evaluated by ability to predict true intake (detect diet-health associations)

- **R-squared** value of the predictor
  - The R-squared value of a predictor is defined as the squared correlation coefficient between true and predicted intake
  - Equivalently, it can be thought of as the proportion of variation in true intake that is explained by the predictor

R-squared value is a direct measure of the ability to predict true intake

- Variance (precision) of estimated diet-health association
- **Power** to detect the association
- Sample size needed to obtain desired power

Comparisons based on data from the Eating at America’s Table Study (EATS)

- Conducted 1997-1998
- Representative sampling of U.S. population
- 965 men and women, aged 20-70

Dietary instruments:

- Four 24HR
  - Administered 3 months apart
  - By telephone
  - Multiple-pass methodology (USDA)
- One FFQ
  - Diet History Questionnaire (NCI)

Dietary variables:

- Total fat
- Whole grains
- Dark-green vegetables

Dietary variables are energy-adjusted (residual method)

Carroll et al. looked at 10 dietary components, both unadjusted for energy and energy-adjusted
Assumptions:

- 24HR provides an unbiased estimate of true usual intake for each individual
- 24HR and FFQ have non-differential error with respect to health outcome

Comparing study designs

- Comparison does not explicitly consider cost of study designs (will be discussed later)
- To simplify comparison, will ignore uncertainty due to estimating parameters in the prediction equation

Comparing different combinations

Assumptions:

- Comparison does not explicitly consider cost of study designs (will be discussed later)
- To simplify comparison, will ignore uncertainty due to estimating parameters in the prediction equation

Comparing different combinations

Research questions:

- Research questions:
  1. Does a single FFQ work better or worse than (one or more) 24HR?
  2. How many 24HR per subject?
  3. How much does adding the FFQ improve the performance of the 24HR (and vice versa)?
  4. Is it better to add another 24HR or add the FFQ?

Comparing different combinations

Graphical comparisons

- Three ways of looking at the data:
  - R-squared value
  - Power to detect diet-health associations
  - Sample size needed to achieve 90% power
- Comparisons relative to the “best” predictor = FFQ plus twelve 24HR
- Results presented for women (results for men are similar)
Combining self-report dietary assessment instruments to reduce the effects of measurement error.

### Power to detect association: total fat

**EATS Study (Women)**

- **Power**
  - 24HR Only
  - 24HR + FFQ
  - FFQ Only

### Percentage increase in sample size: total fat

**EATS Study (Women)**

- **Percentage Increase**
  - 24HR Only
  - 24HR + FFQ
  - FFQ Only

### Comparing different combinations

- **Power to detect association: total fat**
  - EATS Study (Women)
- **Percentage increase in sample size: total fat**
  - EATS Study (Women)

### Ratio of R-squared values: whole grains

**EATS Study (Women)**

- **Ratio of R-squared**
  - FFQ Only
  - 24HR Only
  - 24HR + FFQ

### Power to detect association: whole grains

**EATS Study (Women)**

- **Power**
  - 24HR Only
  - 24HR + FFQ
  - FFQ Only

### Percentage increase in sample size: whole grains

**EATS Study (Women)**

- **Percentage Increase**
  - 24HR Only
  - 24HR + FFQ
  - FFQ Only

### Ratio of R-squared values: dark-green vegetables

**EATS Study (Women)**

- **Ratio of R-squared**
  - 24HR Only
  - 24HR + FFQ
  - FFQ Only
Combining self-report dietary assessment instruments to reduce the effects of measurement error

**Power to detect association: dark-green vegetables**

- EATS Study (Women)

**Percentage increase in sample size: dark-green vegetables**

- EATS Study (Women)

---

**Summary of comparisons**

- In general, calibrated FFQ performs about as well as two 24HR
- For some dietary variables (e.g., dark-green veg.) FFQ performs better than 6 or more 24HR
- Using 4-6 24HR seems to capture most of the information available in 24HR

---

**Summary of comparisons**

- Combining FFQ and 24HR can lead to substantial gains over either alone
- Adding an FFQ to a 24HR is usually better than adding a second 24HR
- For episodically-consumed dietary components, it may be especially important to include an FFQ

---

**Limitations and other considerations**

- Had to simulate 5 or more 24HR (assumes quality will not drop off)
- Study designs with FFQ alone or just a single 24HR require a calibration sub-study of participants who complete two 24HR
- Did not take into account the uncertainty due to estimating parameters in prediction equation
- Assumed that the 24HR provided an unbiased estimate of true intake for each individual
Combining self-report dietary assessment instruments to reduce the effects of measurement error

**Limitations of comparisons**

- Studies with reference biomarkers of intake (doubly-labeled water for total energy, urinary nitrogen for protein) have shown that 24HR are biased for these nutrients.
- In general, incorrectly assuming that the 24HR is unbiased leads to:
  - Biased estimates of diet-health associations
  - Invalid comparisons of precision and power, unless bias is the same for all instruments

**OPEN study**

- Use OPEN to examine effect of biased 24HR
- OPEN study (1999-2000)
- 484 men and women, aged 40-69
- Dietary Assessment:
  - FFQ (2 per subject)
  - 24HR (2 per subject)
  - Reference biomarkers for energy, protein and potassium

**Ratio of R-squared values: protein**

OPEN Study (Men)

- **Ratio of R-squared values: potassium**

OPEN Study (Women)

- **Ratio of R-squared values: potassium**

OPEN Study (Men)
Summary of OPEN study

- In 3 out of 4 cases, assuming 24HR is unbiased produces very similar comparisons as reference biomarkers known to be unbiased
- When comparing study designs assuming 24HR is unbiased
  - Conclusions about any particular dietary component may or may not be valid
  - Conclusions about general patterns that are consistent over many dietary components are probably valid

Is it worth the cost?

- Gain in precision vs. cost
- 24HR and FR impose substantial burden on participants
- New automated 24HR/FR reduce cost but not burden

Limitations and other considerations

Other questions

- How many 24HR can we reasonably expect participants to complete?
  - Response rates?
  - Declining quality?
- Will automated 24HR perform as well as the traditional 24HR?
- Will FR perform similarly to 24HR?
  - Does a 4-day FR = four 24HR?

Looking for answers

- Biomarker studies designed to answer these questions (and more)
  - Six ASA24
  - Two FR
  - Two FFQ
  - Biomarkers of energy, protein and potassium
  - Also: ACT24 (physical activity), accelerometers, blood

Repeat FFQ?

- What about using more than one FFQ?
  - Less within-person variation, so less potential for gain in precision
  - Challenges in interpretation:
    Do differences in two FFQ taken 1 year apart reflect random within-person error or a real change in diet?
    How to define true usual intake if diet is changing over time?

SUMMARY
Combining self-report dietary assessment instruments to reduce the effects of measurement error

**Summary**

- Combining self-report dietary instruments can lead to significant improvement in estimating diet-health associations
- Regression calibration is an effective way to combine instruments

**Summary**

- 4-6 24HR capture most of the information available in 24HR
- Adding FFQ to 1 or more 24HR generally improves prediction more than adding another 24HR

**Summary**

- When designing diet-health studies, one should consider using FFQ plus 4-6 24HR to measure diet
- Other factors such as cost and participant burden must also be considered and balanced with need for precision and power

**Summary**

- These conclusions:
  - Apply to estimating diet-health relationships (predicting individual intake)
  - Do not apply to estimating population distributions of dietary intake
- Tooze et al. (*J Am Diet Assoc*, 2006) found that adding FFQ to two 24HR did not improve estimated population distributions

**QUESTIONS & ANSWERS**

Moderator: Amy Subar

Please submit questions using the Chat function

Next Session

Tuesday, November 29, 2011
10:00-11:30 EST

Combining self-report dietary intake data and biomarker data to reduce the effects of measurement error

Laurence Freedman
Gertner Institute